



IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Application No.: 09/868,987

Applicant: MURDIN, Andrew D. et al.

Filed: December 23, 1999

TC/A.U.: 1645

Examiner: Padmavathi Baskar

Docket No: 032931/0253

DECLARATION PURSUANT TO 37 CFR 1.131

I, Andrew Murdin, Director, External R&D Canada, Aventis Pasteur, hereby declare that:

1. I am an inventor in the above-identified application, and am employed by the assignee, Aventis Pasteur.

2. Details of my employment history are as follows:

Since 2002 Director, External R&D Canada, Aventis Pasteur.

1999-2002 Principal Research Scientist, Aventis Pasteur.

1997-2002 Section Head, Aventis Pasteur.

1993-2003 Project Leader (Chlamydia), Aventis Pasteur.

1990-1993 Research Scientist, Connaught Laboratories Ltd. (subsequently Pasteur Merieux Connaught, subsequently Aventis Pasteur), Toronto, Canada.

1988-1990 Post-Doctoral Research Associate, Dept. of Microbiology, State University of New York, Stony Brook, NY, USA.

1985-1987 Post-Doctoral Research Fellow, Dept. of Microbiology, University of Surrey, Guildford, Surrey, England.

1981-1985 Scientific Officer, Vaccine Research Dept., Animal Virus Research Institute, Pirbright, Surrey, England.

Details of my education are as follows:

B.Sc., University of Bath, England, 1980

Ph.D., University of Surrey, England, 1986.

3. Attached is a copy of a document signed by me and Dathao Ho, an employee of the assignee Aventis Pasteur. The document sets forth the amino acid sequence, the nucleotide sequence, and the restriction map of CPN100686.

4. CPN100686 corresponds to the gene identified as SEQ ID Nos: 1 and 14 of the application. The amino acid sequence of CPN100686 set forth in the attached document is identical to SEQ ID No:14 of this application. The sequences and restriction map set forth in the document are essentially the same as those shown in Figures 1 and 14 of this application.

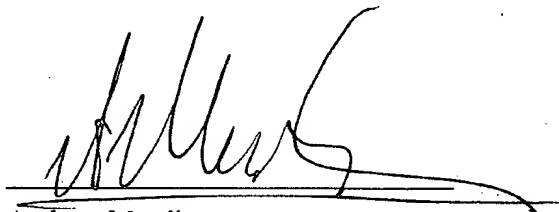
5. The dates deleted from the attached document are before November 4, 1998.

6. The document shows that I possessed these amino acid and nucleotide sequences prior to November 4, 1998.

7. I hereby declare that all statements made herein of my knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like are punishable by fine or imprisonment, or both, under section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of this application or any patent issuing thereon.

Mord 6ⁿ 2504

Date



Andrew Murdin

Director, External R&D Canada, Aventis Pasteur

CONFIDENTIAL

>CPN100686 protein-export membrane protein SecD
VSSPILNVPLKNHASVSGKFTHREVSKLASDLKSGAMSFVPEVLSEETISSDLGKKQCTQGIISACCGLAMLIVL
MSVYYRFGGVIASGAVLLNLLLIWAALQYLDAPLTLSGLAGIVLAMGMAVDANVLVFERIREEFLLSQSLKKSVE
KGYTKAFGAIFDSNLT'TVLASALLFFLDTGPIKGFALTLILGIFSSMFTALFMTKFFFMLWMNKTQHTQLHMMNK
FVGIKHDFLRGCKKLWAVSGSVFLLGCVLGFAGWNSVLGMDFKGGYAFTFNPKEHGISDVAQMRGKVHKLQEA
GLSSRDFRIQTFGSSEKIKIYFSDKSFKLYXSRYEPLSXNXRSXAGVNVGLLSETGLDFSTETLNETQNFWSKVS
SKLSKKMRYQATIGLLGALAIILLYVSLRFEWQYAFSAVCALIHDLATCAVLFI AHFFLKKIQIDLQAIGALMT
VLGYSLNNTLIIFDRIREDRQANLFTPMHVLVNDALQKTF SRTVMTTATTL SVLLMLLFIGGSSVFNFAPIMTIG
ILLGTLSSLYIAPPLL

ACTUAL ENCODED SEQUENCE

1 MVSSPILNVP LKNHASVSGK FTHREVSKLA SDLKSGAMSF VPEVLSEETI
51 SSDLGKKQCT QCIISACCCL AMLIVLMSVY YRFGGVIASG AVLLNLLLIW
101 AALQYLDAPL TLSGLAGIVL AMGMAVDANV LVFERIREEF LLSQSLKKS
151 EKGYTKAFGA IFDSNLT'TVL ASALLFFLD TGPIKGFALTL ILGIFSSMFT
201 ALFMTKFFFM LWMNKTQHTQ LHMMNKFVGI KHDFLRGCKK LWAVSGSVFL
251 LGCVALGFAG WNSVLGMDFK GGYAFTFNPKEHGISDVAQM RGKVHKLQEA
301 AGLSSRDFRI QTFGSSEKIK IYFSDKALS TKQIRASLLK LTIMSWRYCG
351 IVVRNRPRFL YGNSKRNAKF WSKVSSKLSK KMRYQATIGL LGALAIILLY
401 VSLRFEWQYA FSAVCALIHDLATCAVLFI AHFFLKKIQIDLQAIGALMT
451 VLGYSLNNTLIIFDRIREDRQANLFTPMHV LVNDALQKTF SRTVMTTATTL
501 LSVLLMLLFI GGSSVFNFAPIMTIGILLGT LSSLYIAPPL LLFMVRKENR
551 SK*

CODING SEQUENCE

THE PROTEIN IS ENCODED ON THE POSITIVE STRAND

The ATG is presumably the start codonThe TAA is presumably the stop codon

```
1  ATGGACTTCC GCATATTGTC AGGAGGGGAT CAGCGGCACT GCTAATGGAC
51  AATATTCTGC AAACCGTGGA TGGCGTATGG CTGTAGTGAT TGACGGTTAT
101 ATGGTCAGCA GCCCTATTTT AAACGTCCCA TTGAAAAATC ATGCCAGTGT
151 CTCAGGGAAA TTTACCCACC GTGAAGTGAG CAAACTCGCC TCAGATTTAA
201 AATCTGGAGC GATGTCTTTT GTTCCCAGAG TTCTCAGTGA AGAGACGATC
251 TCTTCTGATC TTGGGAAAAA ACAATGTACA CAAGGCATTA TCTCAGCATG
301 CTGTGGCTTG GCAATGCTTA TTGTTTTGAT GAGCGTATAT TATAGATTTG
351 GAGGCGTCAT CGCTTCGGGA GCTGTTCTTC TGAATCTTTT GCTTATCTGG
401 GCAGCTCTAC AGTATTTGGA TGCGCCACTC ACCTTGTGAG GACTCGCTGG
451 GATTGTCTCTT GCTATGGGGA TGGCCGTAGA TGCAAATGTT CTTGTATTCG
501 AAAGAATCCG AGAGGAATTT TTATTGTCTC AAAGTCTTAA AAAATCTGTA
551 GAAAAAGGAT ATACCAAGGC TTTTGGAGCC ATTTTGGATT CTAAGTTGAC
601 TACAGTATTG GCCTCAGCAC TTCTTTTCTT CCTAGATACA GGGCCTATTA
651 AAGGGTTTGC TTTGACATTG ATTTTAGGAA TTTTCTCTTC AATGTTTACG
701 GCTCTTTTCA TGAATAAAT TTTCTTCATG CTGTGGATGA ATAAGACCCA
751 ACATACACAG TTGCATATGA TGAATAAGTT CGTGGGGATA AAGCATGATT
801 TCTTGAGAGG ATGCAAAAAA CTTTGGGCTG TTTCTGGAAG TGT'TTTTCTT
851 TTAGGTTGCG TTGCTCTCGG GTTTGGAGCC TGGGAATCCG TTTTGGGAAT
901 GGATTTTAAA GGAGGGTATG CCTTTACCTT TAATCCAAAA GAGCATGGCA
951 TCAGCGATGT TGCTCAAATG CGTGGCAAAG TTGTGCATAA ACTACAGGAA
1001 GCTGGTCTTT CTTCTAGAGA CTTCCGTATT CAAACATTTG GATCTTTCAGA
1051 AAAGATCAAA ATCTATTTTA GTGATAAAGC TTTAAGCTAT ACTAAGCAGA
1101 TACGAGCCTC TCTCCTAAAA TTAACGATCA TGAGCTGGCG TTATTGTGGG
1151 ATTGTTGTCA GAAACAGGCC TAGATTTCTC TACGGAAACT CTAAACGAAA
1201 CGCAAAATTT TGGTCAAAGG TAAGCAGCAA ACTATCGAAG AAAATGCGTT
1251 ATCAGGCGAC CATCGGGCTT TTAGGAGCTT TGGCAATCAT CTTGCTCTAT
1301 GTGAGTTTGC GCTTTGAATG GCAATATGCT TTCAGTGCCG TATGCGCTTT
1351 AATTCATGAC CTTTGGGCTA CCTGTGCAGT CTTGTTTATA GCACATTTCT
1401 TTTTGAAGAA AATTCAAATA GATTTGCAAG CCATTGGTGC TTTAATGACT
1451 GTATTGGGGT ATTCATTAAA CAATACTTTG ATCATTTTTG ATCGTATTCG
```

1501 TGAAGATCGC CAAGCGAACC TGTTTACCCC TATGCATGTT TTAGTTAATG
1551 ATGCCCTTCA AAAGACGTTC AGCCGCACGG TAATGACAAC AGCTACAAC
1601 CTATCAGTTT TGTTAATGCT TTTGTTTATA GCGGGCTCCT CTGTCTTTAA
1651 TTTTGCATTT ATTATGACCA TAGGGATTCT TCTAGGAACT TTATCGTCTC
1701 TTTATATTGC ACCACCTCTG TTGTTGTTTA TGGTCCGTAA AGAAAATCGC
1751 TCAAAATAAG TACCGTTAAA CTTAATCTAA CGTGTAGCAA TATAAAAATC
1801 TCCTTTGGGA CTTTAGTCCC AAAGGCCCT GTGGTATTAA ATTTATGACA
1851 AATTCAGATA ATGC

SEQUENCE ALIGNMENT

```

101 ATGGTCAGCAGCCCTATTTTAAACGTCCCATTGAAAAATCATGCCAGTGT 150
    |||
    1 MetValSerSerProIleLeuAsnValProLeuLysAsnHisAlaSerVa 17
151 CTCAGGGAAATTTACCCACCGTGAAGTGAGCAAACCTCGCCTCAGATTTAA 200
    |||
    18 lSerGlyLysPheThrHisArgGluValSerLysLeuAlaSerAspLeuL 34
201 AATCTGAGCGATGTCTTTTGTTCCTCCGAGGTTCTCAGTGAAGAGACGATC 250
    |||
    35 ysSerGlyAlaMetSerPheValProGluValLeuSerGluGluThrIle 50
251 TCTTCTGATCTTGGGAAAAACAATGTACACAAGGCATTATCTCAGCATG 300
    |||
    51 SerSerAspLeuGlyLysLysGlnCysThrGlnGlyIleIleSerAlaCy 67
301 CTGTGGCTTGGCAATGCTTATTTGTTTGGATGAGCGTATATTATAGATTG 350
    |||
    68 sCysGlyLeuAlaMetLeuIleValLeuMetSerValTyrTyrArgPheG 84
351 GAGGCGTCATCGCTTCGGGAGCTGTTCTTCTGAATCTTTTGCTTATCTGG 400
    |||
    85 lyGlyValIleAlaSerGlyAlaValLeuLeuAsnLeuLeuLeuIleTrp 100
401 GCAGCTCTACAGTATTTGGATGCGCCACTCACCTTGTCAGGACTCGCTGG 450
    |||
    101 AlaAlaLeuGlnTyrLeuAspAlaProLeuThrLeuSerGlyLeuAlaGl 117
451 GATTGTTCTTGCTATGGGGATGGCCGTAGATGCAATGTTCTTGTATTCTG 500
    |||
    118 yIleValLeuAlaMetGlyMetAlaValAspAlaAsnValLeuValPheG 134
501 AAAGAATCCGAGAGGAATTTTATTGTCTCAAAGTCTTAAAAAATCTGTA 550
    |||
    135 luArgIleArgGluGluPheLeuLeuSerGlnSerLeuLysLysSerVal 150
551 GAAAAAGGATATACCAAGGCTTTTGGAGCCATTTTGGATTCTAACTTGAC 600
    |||
    151 GluLysGlyTyrThrLysAlaPheGlyAlaIlePheAspSerAsnLeuTh 167
601 TACAGTATTGGCCTCAGCACTTCTTTTCTCTAGATACAGGGCCTATTA 650
    |||
    168 rThrValLeuAlaSerAlaLeuLeuPhePheLeuAspThrGlyProIleL 184
651 AAGGGTTTGCTTTGACATTGATTTTAGGAATTTCTCTTCAATGTTTACG 700
    |||
    185 ysGlyPheAlaLeuThrLeuIleLeuGlyIlePheSerSerMetPheThr 200
701 GCTCTTTTCATGACTAAATTTTCTTCATGCTGTGGATGAATAAGACCCA 750
    |||
    201 AlaLeuPheMetThrLysPhePhePheMetLeuTrpMetAsnLysThrGl 217
751 ACATACACAGTTGCATATGATGAATAAGTTCGTGGGATAAAGCATGATT 800
    |||
    218 nHisThrGlnLeuHisMetMetAsnLysPheValGlyIleLysHisAspP 234
801 TCTTGAGAGGATGCAAAAACTTTGGGCTGTTTCTGGAAGTGTTTTCTT 850
    |||
    235 heLeuArgGlyCysLysLysLeuTrpAlaValSerGlySerValPheLeu 250

```

851	TTAGGTTGCGTTGCTCTCGGGTTTGGAGCCTGGAATCCGTTTGGGGAAT	900
251	LeuGlyCysValAlaLeuGlyPheGlyAlaTrpAsnSerValLeuGlyMe	267
901	GGATTTTAAAGGAGGGTATGCCTTTACCTTTAATCCAAAAGAGCATGGCA	950
268	tAspPheLysGlyGlyTyrAlaPheThrPheAsnProLysGluHisGlyI	284
951	TCACGCGATGTTGCTCAAATGCGTGGCAAAGTTGTGCATAAACTACAGGAA	1000
285	leSerAspValAlaGlnMetArgGlyLysValValHisLysLeuGlnGlu	300
1001	GCTGGTCTTTCTTCTAGAGACTTCCGTATTCAAACATTGGATCTTCAGA	1050
301	AlaGlyLeuSerSerArgAspPheArgIleGlnThrPheGlySerSerGl	317
1051	AAAGATCAAAATCTATTTTAGTGATAAAGCTTTAAGCTATACTAAGCAGA	1100
318	uLysIleLysIleTyrPheSerAspLysAlaLeuSerTyrThrLysGlnI	334
1101	TACGAGCCTCTCTCCTAAATTAACGATCATGAGCTGGCGTTATTGTGGG	1150
335	leArgAlaSerLeuLeuLysLeuThrIleMetSerTrpArgTyrCysGly	350
1151	ATTGTTGTCAGAAAAGGCCCTAGATTTCTCTACGGAACTCTAAACGAAA	1200
351	IleValValArgAsnArgProArgPheLeuTyrGlyAsnSerLysArgAs	367
1201	CGCAAAATTTTGGTCAAAGGTAAGCAGCAAACATATCGAAGAAAATGCGTT	1250
368	nAlaLysPheTrpSerLysValSerSerLysLeuSerLysLysMetArgT	384
1251	ATCAGGCGACCATCGGGCTTTTAGGAGCTTTGGCAATCATCTTGCTCTAT	1300
385	yrGlnAlaThrIleGlyLeuLeuGlyAlaLeuAlaIleIleLeuLeuTyr	400
1301	GTGAGTTTGCGCTTTGAATGGCAATATGCTTTCAGTGCCGTATGCGCTTT	1350
401	ValSerLeuArgPheGluTrpGlnTyrAlaPheSerAlaValCysAlaLe	417
1351	AATTCATGACCTTTTGGCTACCTGTGCAGTCTTGTTTATAGCACATTCT	1400
418	uIleHisAspLeuLeuAlaThrCysAlaValLeuPheIleAlaHisPheP	434
1401	TTTTGAAGAAAATTCAAATAGATTTGCAAGCCATTGGTGCTTTAATGACT	1450
435	heLeuLysLysIleGlnIleAspLeuGlnAlaIleGlyAlaLeuMetThr	450
1451	GTATTGGGGTATTTCATTAAACAATACTTTGATCATTTTGGATCGTATTTCG	1500
451	ValLeuGlyTyrSerLeuAsnAsnThrLeuIleIlePheAspArgIleAr	467
1501	TGAAGATCGCCAAGCGAACCTGTTTACCCCTATGCATGTTTATAGTTAATG	1550
468	gGluAspArgGlnAlaAsnLeuPheThrProMetHisValLeuValAsnA	484
1551	ATGCCCTTCAAAGACGTTACAGCCGCACGGTAATGACAACAGCTACAAC	1600
485	spAlaLeuGlnLysThrPheSerArgThrValMetThrThrAlaThrThr	500
1601	CTATCAGTTTGTGTTAATGCTTTTGTGTTATAGGCGGCTCCGTCTGTCTTAA	1650
501	LeuSerValLeuLeuMetLeuLeuPheIleGlyGlySerSerValPheAs	517

```

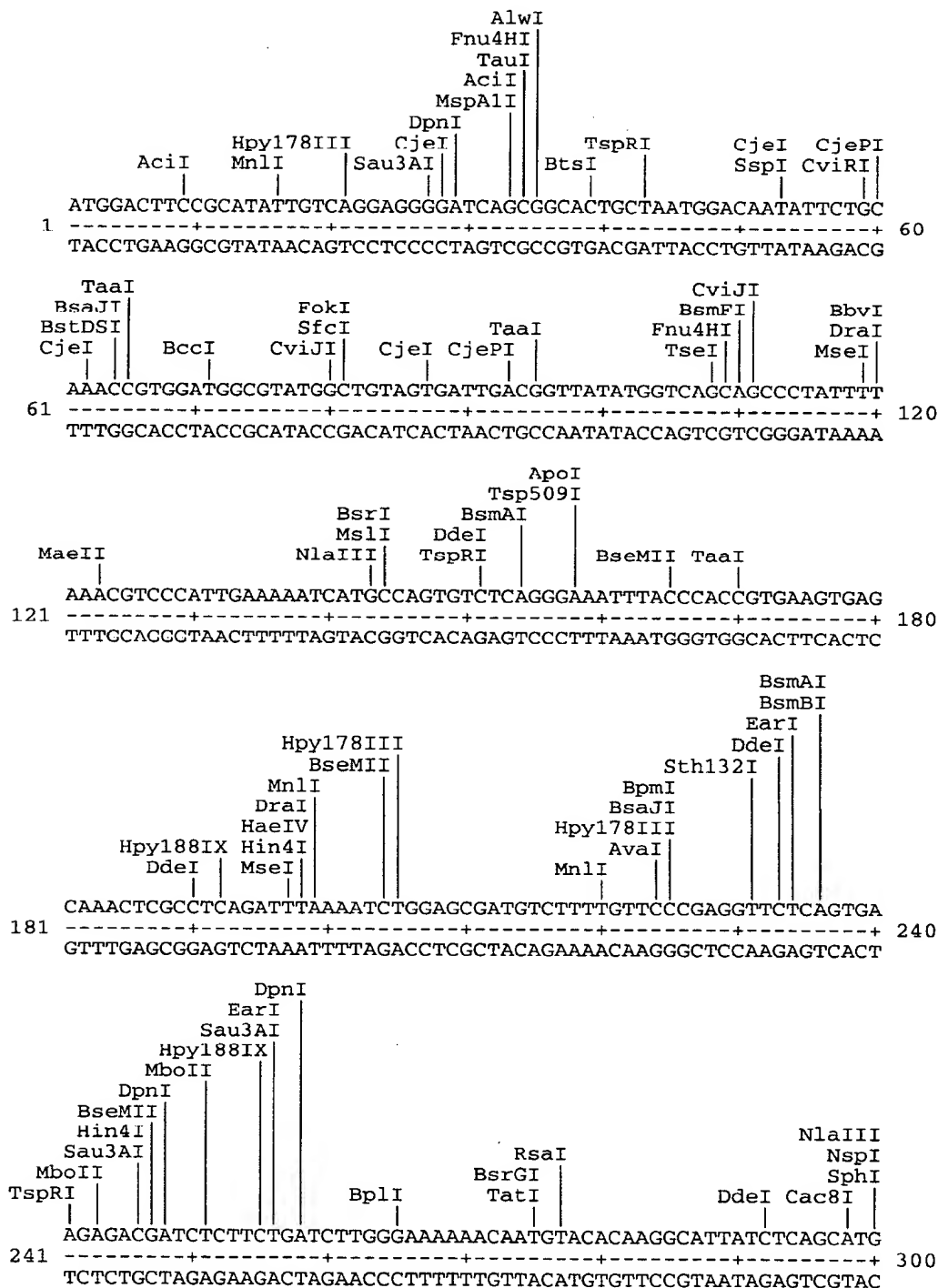
1651 TTTTGCATTTATTATGACCATAGGGATTCTTCTAGGAAC TTTATCGTCTC 1700
      |||||||||||||||||||||||||||||||||||||||
518  nPheAlaPheIleMetThrIleGlyIleLeuLeuGlyThrLeuSerSerL 534

1701 TTTATATTGCACCACCTCTGTTGTTGTTTATGGTCCGTAAAGAAAATCGC 1750
      |||||||||||||||||||||||||||||||||||||||
535  euTyrIleAlaProProLeuLeuLeuPheMetValArgLysGluAsnArg 550

1751 TCAAAA 1756
      |||||
551  SerLys 552

```


RESTRICTION MAP



BseMII CviJI BsrDI MnlI HgaI AceIII Sth132I Hin4I BsaHI
 301 CTGTGGCTTGGCAATGCTTATTGTTTTGATGAGCGTATATTATAGATTGGAGGCGTCAT 360
 GACACCGAACCGTTACGAATAACAAACTACTCGCATATAATATCTAAACCTCCGCAGTA

AluI CviJI MboII MwoI Hpy178III HinfI TfiI Hpy188IX Fnu4HI TseI CjeI AluI CviJI SfaNI SfcI AceIII BbvI TaaI
 361 CGCTTCGGGAGCTGTTCTTCTGAATCTTTTGCTTATCTGGGCAGCTCTACAGTATTTGGA 420
 GCGAAGCCCTCGACAAGAAGACTTAGAAAACGAATAGACCCGTCGAGATGTCATAAACCT

HphI HhaI FokI CjeI Hpy178III HinfI PleI BceFI SfaNI GdIII EaeI BccI HaeIII CviJI MwoI
 421 TCGCCCACTCACCTTGTCAGGACTCGCTGGGATTGTTCTTGCTATGGGGATGGCCGTAGA 480
 ACGCGGTGAGTGGAACAGTCTGAGCGACCCTAACAAAGACGATACCCCTACCGGCATCT

CviRI FokI NspV HinfI Hpy188IX MnlI ApoI Tsp509I BsmAI MseI
 481 TGCAAATGTTCTTGTATTTCGAAAGAATCCGAGAGGAATTTTATTGTCTCAAAGTCTTAA 540
 ACGTTTACAAGAACATAAGCTTTCTTAGGCTCTCTTAAAAATAACAGAGTTTCAGAATT

SfcI BsaJI StyI CviJI NlaIV MwoI HinfI TfiI SfcI
 541 AAAATCTGTAGAAAAAGGATATACCAAGGCTTTTGGAGCCATTTTGTATTCTAACTTGAC 600
 TTTTAGACATCTTTTCTTATATGGTTCCGAAAACCTCGGTAAAACTAAGATTGAACTG

BbvCI Bpu10I DdeI CviJI HaeI MnlI BseMII BfaI Eco0109I Sau96I CviJI HaeIII BslI MseI
 601 TACAGTATTGGCCTCAGCACTTCTTTTCTTCCTAGATACAGGGCTATTAAAGGGTTTGC 660
 ATGTCATAACCGGAGTCGTGAAGAAAAGAAGGATCTATGTCCCGGATAATTTCCTCAAACG

Diagram illustrating the restriction enzyme sites for the *hcr* gene, showing the DNA sequence and the positions of various restriction enzymes.

The DNA sequence is shown in 100 bp fragments, with restriction enzyme sites indicated by vertical lines and labels above the sequence. The sequence is as follows:

661 TTTGACATTGATTTTAGGAATTTTCTCTTCAATGTTACGGCTCTTTTCATGACTAAATT 720
-----+-----+-----+-----+-----+-----+-----+-----+
AAACTGTAACATAAAATCCTTAAAAGAGAAGTTACAAATGCCGAGAAAAGTACTGATTTAA

721 TTTCTTCATGCTGTGGATGAATAAGACCCAACATACACAGTTGCATATCATCAATAAGTT 780
-----+-----+-----+-----+-----+-----+-----+-----+
AAAGAAGTACGACACCTACTTATTCTGGGTTGTATGTGTCAACGTATACTACTTATTCAA

781 CGTGGGGATAAAGCATGATTTCTTGAGAGGATGCAAAAAACTTTGGGCTGTTTCTGGAAG 840
-----+-----+-----+-----+-----+-----+-----+-----+
GCACCCCTATTTCGTACTAAAGAACTCTCCTACGTTTTTTTGAACCCGACAAAGACCTTC

841 TGTTTTTCTTTTAGGTTGCGTTGCTCTCGGGTTTGGAGCCTGGAATTCCGTTTGGGAAT 900
-----+-----+-----+-----+-----+-----+-----+-----+
ACAAAAAGAAAATCCAACGCACGAGAGCCCAAACCTCGGACCTTAAGGCAAAACCCTTA

901 GGATTTTAAAGGAGGGTATGCCTTTACCTTTAATCCAAAAGAGCATGGCATCAGCGATGT 960
-----+-----+-----+-----+-----+-----+-----+-----+
CCTAAAAATTTCTCTCCATACGGAAATGGAATTAGGTTTTCTCGTACCGTAGTCGCTACA

961 TGCTCAAATGCGTGGCAAAGTTGTGCATAAACTACAGGAAGCTGGTCTTTCTTCTAGAGA 1020
-----+-----+-----+-----+-----+-----+-----+-----+
ACGAGTTTACGCACCGTTTCAACACGTATTTGATGTCCTTCGACCAGAAAGAAGATCTCT

Restriction enzyme sites (indicated by vertical lines and labels):

- ApoI, MboII, Tsp509I, EarI, CviJI, Hpy178III, NlaIII, BceFI, BsaBI, DpnI, Sau3AI, AlwI, Hpy188IX, NdeI, CviRI, TaaI, XmnI, SmlI, MnlI, SfaNI, NlaIII, CviRI, Bce83I, FokI, Hpy178III, ApoI, EcoRI, Tsp509I, ScrFI, CviJI, EcoRII, NlaIV, Sth132I, AvaI, DraI, MseI, MnlI, MseI, NlaIII, SfaNI, MboII, AluI, CviJI, BsmAI, BfaI, XbaI, Hpy178III, CviRI, SfcI, BsaBI, DpnI, Sau3AI, AlwI, Hpy188IX.

Tth111III
 DpnI
 BstYI
 Sau3AI
 Eco57I MboII
 AluI
 CviJI
 HindIII
 1021 CTTCCGTATTCAAACATTTGGATCTTCAGAAAAGATCAAAATCTATTTTAGTGATAAAGC 1080
 GAAGGCATAAGTTTGTAAACCTAGAAGTCTTTTCTAGTTTGTAGATAAAATCACTATTTTCG
 Cac8I
 RleAI
 AluI
 CviJI
 NlaIII
 Hpy178III
 RcaI
 BplI
 DpnI
 Sau3AI
 MseI
 AceIII
 Tsp509I
 MnlI
 AluI
 CviJI
 MseI
 DdeI
 Hin4I
 CviJI
 1081 TTTAAGCTATACTAAGCAGATACGAGCCTCTCTCTAAATTAACGATCATGAGCTGGCG 1140
 AAATTCGATATGATTTCGTCTATGCTCGGAGAGAGGATTTTAATTGCTAGTACTCGACCGC
 BfaI
 CviJI
 HaeI
 HaeIII
 StuI
 Hpy188IX
 1141 TTATTGTGGGATTGTTGTCAGAAACAGGCCTAGATTTCTCTACGGAAACTCTAAACGAAA 1200
 AATTACACCCTAACACAGTCTTTGTCCGGATCTAAAGAGATGCCTTTGAGATTTGCTTT
 BcgI
 Fnu4HI
 TseI
 BbvI
 TaqI
 Sth132I
 MboII
 BcgI
 1201 CGCAAAATTTTGGTCAAAGGTAAGCAGCAAACTATCGAAGAAAATGCGTTATCAGGCGAC 1260
 GCGTTTTTAAACAGTTTCCATTCGTCGTTTGATAGCTTCTTTTACGCAATAGTCCGCTG
 AluI
 CviJI
 BccI
 HhaI
 1261 CATCGGGCTTTTAGGAGCTTTGGCAATCATCTTGCTCTATGTGAGTTTGCGCTTTGAATG 1320
 GTAGCCCGAAAATCCTCGAAACCGTTAGTAGAACGAGATACACTCAAACGCGAAACTTAC
 NlaIII
 Hpy178III
 Tsp509I
 MseI
 BceI
 MwoI
 TspRI
 HhaI
 MwoI
 RcaI
 CviJI
 MwoI
 CviRI
 1321 GCAATATGCTTTTCAGTGCCGATGCGCTTTAATTCATGACCTTTTGGCTACCTGTGCAGT 1380
 CGTTATACGAAAGTCACGGCATACGCGAAATTAAGTACTGGAAAACCGATGGACACGTCA

BsgI ApoI CviJI
 Tsp509I MboII CviRI Cac8I MwoI
 1381 CTTGTTTATAGCACATTTCTTTTGAAGAAAATTCAAATAGATTGCAAGCCATTGGTGC 1440
 GAACAAATATCGTGTAAGAAAACTTCTTTTAAGTTTATCTAAACGTTTCGGTAACCACG
 DpnI
 BclI DpnI
 MseI TaaI MseI Sau3AI Sau3AI Hpy178III
 1441 TTTAATGACTGTATTGGGGTATTCATTAAACAATACTTTGATCATTTTGTATCGTATTTCG 1500
 AAATTACTGACATAACCCATAAGTAATTTGTTATGAAACTAGTAAAACTAGCATAAGC
 SfaNI
 NlaIII
 NspI
 NsiI
 DpnI MboII CviRI MseI
 Sau3AI
 1501 TGAAGATCGCCAAGCGAACCTGTTTACCCCTATGCATGTTTGTAGTTAATGATGCCCTTCA 1560
 ACTTCTAGCGGTTTCGCTTGGACAAATGGGGATACGTACAAAATCAATTACTACGGGAAGT
 AciI
 Fnu4HI
 TauI MslI AluI
 MaeII CviJI TaaI CviJI MseI
 1561 AAAGACGTTTCAGCCGCACGGTAATGACAACAGCTACAACCTCTATCAGTTTGTTAATGCT 1620
 TTTCTGCAAGTCGGCGTGCCATTACTGTTGTTCGATGTTGAGATAGTCAAAACAATTACGA
 NlaIV
 CviJI
 Fnu4HI
 TauI MnlI
 BseRI AciI Tsp509I MseI CviRI CjePI MboII HinfI
 TfiI
 1621 TTTGTTTATAGGCGGCTCCTCTGTCTTTAATTTGCATTTATTATGACCATAGGGATTCT 1680
 AAACAAATATCCGCCGAGGAGACAGAAATTAAAACGTAAATTAATAC'TGGTAT'CCCTAAGA
 BsmAI
 BsmBI CviRI MnlI AvaII
 BfaI CjePI Sau96I
 1681 TCTAGGAACCTTATCGTCTCTTTATATTGCACCACCTCTGTTGTTGTTTATGGTCCGTAA 1740
 AGATCCTTGAAATAGCAGAGAAATATAACGTGGTGGAGACAACAACAATACCAGGCATT
 MseI
 TaaI AflIII
 RsaI MseI MaeII
 1741 AGAAAAATCGCTCAAAATAAGTACCGTTAAACTTAATCTAACGTGTAGCAATATAAAAAATC 1800
 TCTTTTAGCGAGTTTTATTTCATGGCAATTTGAATTAGATTGCACATCGTTATATTTTAG
 NlaIV
 CviJI
 HaeIII

>SW:SECF_MYCTU Q50635 mycobacterium tuberculosis. protein-export membrane
protein secf. 11/97
Length = 442

Score = 90 (42.3 bits), Expect = 5.9e-16, Sum P(5) = 5.9e-16
Identities = 30/118 (25%), Positives = 50/118 (42%)

```
Query:   316 SEKIKIYFSDKALSYTKQIRASLLKLTIMSWRXCGIVVRNRPRFLYGNSKRNAKFWSKVS 375
          SE  +      S T QIR+ L      +      +      P+      G      + A      S VS
Sbjct:   121 SEPQSVVIVGAGASATVQIRSETLTSDQTAKLRDALFEAFGPKGTDGQPSKQAISDSAVS 180

Query:   376 SKLSKKMRYQATIGLLGALAIILLYVSLRFQYAFSAVCALIHDLATCAVLFIHF 433
          ++  +A I L+  L ++ LY+++R+E      SA+ A++ DL  T  V  +  F
Sbjct:   181 ETWGGQITKKAIVLVFLVLVALYITVRYERYMTISAITAMLFDLTVTAGVYSLVGF 238
```